The Optimization of Framed Aloha based RFID Algorithms

Lei Zhu
The Chinese University of Hong Kong
Shatin, New Territory
Hong Kong, China
lzhu6@ie.cuhk.edu.hk

Tak-Shing Peter Yum
The Chinese University of Hong Kong
Shatin, New Territory
Hong Kong, China
tsyum@ie.cuhk.edu.hk

ABSTRACT

The anti-collision mechanism is a very important part in Radio-frequency Identification (RFID) systems. Among all the algorithms, the Framed Aloha based (FA) ones are most widely used due to simplicity and robustness. Previous works mainly focused on the tag population estimation, but determined the reading strategy based on the classical results of Random Access (RA) systems. We show that a new theory is needed for the optimization of the RFID systems as they have characteristics very different from the RA systems. In this paper, We propose a new approach to minimize the total expected reading time by choosing the most suitable frame size based on the tag population distribution. We show that the optimal strategy can be used in different applications. The mathematical analysis and computer simulation show our approach outperforms the previous optimization works in the literature.

Categories and Subject Descriptors

F.2.2 [**Analysis of Algorithms and Problem Complexity**]: Nonnumerical Algorithms and Problems; G.3 [**Probability and Statistics**]: Distribution functions, Markov processes

General Terms

Algorithms, Design, Performance, Theory

Keywords

Algorithm, RFID, Framed Aloha, Optimization

1. INTRODUCTION

In Radio-frequency Identification (RFID) systems, tags share a common communication channel. Therefore, if multiple tags transmit at the same time, their packets will collide and get lost [1]. Passive tags have bare-bone functionality and no embedded power supply. They cannot sense the media or cooperate with one another. The RFID reader needs to coordinate their transmissions to avoid collisions. The communication time between tags and readers are slotted. But unlike CSMA system, tags need to **'reserve'** the channel before transmission. The total communication time therefore includes the *contention time* and the *operation time*. To illustrate, consider the tag reading operation in EPCglobal standards [2] as shown in Figure 1. Within a *contention* slot, the reader broadcasts a trigger command ('Query' or 'QueryRep'). After receiving this command, each tag runs a random function to decide whether to reply or not. Tags only reply a short packet named 'RN16' (random number 16 bits). It is used as the temporary ID for this tag. If multiple tags reply or no tag replies, the reader sends an trigger command again. If only one tag replies, the reader can receive the packet successfully and *operate* on this tag by its RN16 after this slot. The operation may include reading data, writing new data, changing password, etc. Since the *operation* slot is collision-free, the total *operation time* does not depend on the reading strategy. Therefore, the performance of anti-collision algorithms is conventionally evaluated by the average contention time measured by the number of contention slots. In literature, the *'slot'* usually refers to the *contention slot* while the *'reading time'* usually refers to the *contention time* measured in contention slots. We follow this convention in this paper.

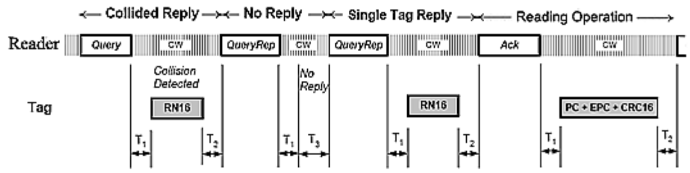

Figure 1: The communication between the reader and tags

Depending on working principles, RFID anti-collision algorithms can be divided into three main types: Tree based algorithms [5][6], Framed Aloha based (FA) algorithms [8-16] and Interval based algorithms [7]. Different types usually require different hardware and software design of both the reader and the tags. Among all the types, FA algorithms are most widely used in RFID communication standards [2-4] due to simplicity and robustness.

In most applications, the number of tags are unknown before identification. So a proper FA algorithm always contains two parts: **Population Estimation** part and **Reading Strat-**

MSWiM'09, October 26–29, 2009, Tenerife, Canary Islands, Spain.

egy Determination part. The first part is for estimating the tag population based on tags' replies while the second part is for adjusting the command parameters, such as the frame size, based on the estimation. Previous works [8-16] emphasized Population Estimation methods and designed reading strategies based on the classic results of Random Access (RA) Systems [10]. Since RFID systems and RA systems are fundamentally different (the details are in section 3), the use of RA results will not lead to the optimal reading strategy in RFID systems. In this paper, we model the reading process as a Markov Chain and derive the optimal reading strategy through first-passage-time analysis. We show that the optimal strategy can be easily incorporated into the different applications to give significant performance improvement, especially when the variance of tag population is large.

In section 2, we introduce the basic ideas of the FA algorithms. In section 3, we give a survey of the traditional strategies of FA algorithms and point out an unjustified assumption used in previous attempts of reading strategy optimization. In section 4, a new model is proposed to derive the optimal reading strategy. In section 5, we show the applications of the optimal strategy and compare its performance with the previous works.

2. FRAMED ALOHA-BASED RFID SYSTEMS

Framed Aloha (FA) is a variation of slotted Aloha where a terminal is permitted to transmit once per frame. The frame size L is broadcast by the reader at the beginning of every round; each tag randomly chooses a value from 0 to $L-1$ as its transmission delay. In RFID systems, the FA algorithms have some special characteristics.

2.1 Limited Choices of Frame Size

Many RFID systems have limitations on the choice of frame size due to hardware constraints. For example, in EPCglobe standards, it is limited to only 16 choices as 2^Q, where $Q = 0, 1, 2, \ldots, 15$.

2.2 Silence Command

In the original design of FA algorithms, tags do not know their transmission results as there is no feedback from the reader. They will all transmit again in the next round of contention. Readers have difficulty ascertaining the end of the reading process as some tags may suffer collisions again and again (tag starvation problem). This situation was changed by the introduction of the Silence Command[1] in EPCglobe standards. After identifying a tag, the reader will broadcast its ID and ask it to keep silent.

2.3 Reset Command

As in Figure 1, the RFID reader has to broadcast a 'Trigger' command in every time slot because tags need to extract power from the command signal to reply. The reader consequently does not have to wait until the end of a frame to change the reply probability by setting the appropriate frame size. Some designs introduce the *Frame-size Reset* command[2] to cancel a running frame and initiate a new one.

[1]It is also referred to as *Kill* command in literature. In EPCglobe standards, it corresponds to the *Select* command, which have other uses besides silencing a tag.

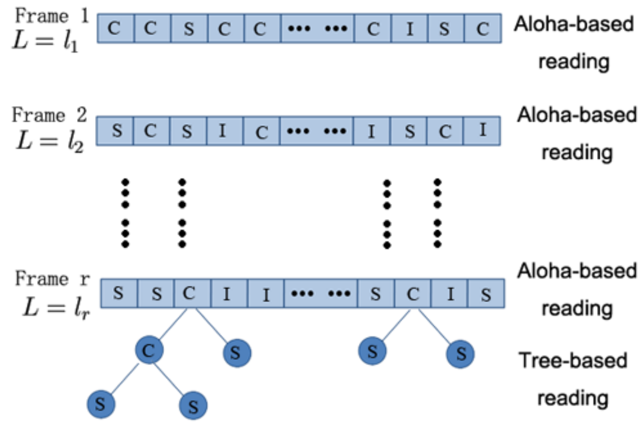

Figure 2: The working mechanism of Type 4 Aloha-based algorithms

2.4 Split Command

Some FA algorithms [8] have the Split Command of Tree-based algorithms embedded. After a frame of reading, the reader may choose to initiate a new frame or just split the collided slots. The reading process of algorithms embedded with the Split Command is illustrated in Figure 2. This was shown to improve the performance at the expense of hardware complexity.

2.5 Classification of RFID Systems

Depending on the tag-reader capability, or the set of commands supported, the RFID systems for FA algorithms can be classified into four types as follows:

1: support only Framed Aloha;

2: support Framed Aloha with Silence Command;

3: support Framed Aloha with Silence and Reset Command;

4: support Framed Aloha with Silence and Split Command.

Note the term *'Type × RFID system'* refers to an RFID system (hardware and software) which supports a certain set of commands while the term *'Type × algorithm'* refers to a reading strategy which determines when and how to use these commands. In this paper, we will focus on the Type 2 RFID system, which is simple enough and relatively efficient.

[2]In EPCglobal standards, it corresponds to the QueryAdjust command.

3. A SURVEY OF PREVIOUS WORKS

FA algorithms are widely used in random-access systems. The classical result for throughput U with N attempting terminals and frame size L is given in [10] as:

$$U(N, L) = \frac{N}{L}\left(1 - \frac{1}{L}\right)^{N-1}. \quad (1)$$

The throughput U can be optimized by setting the frame size equal to the terminal number, or $L = N$. The bound for large N is $U = e^{-1}$. However, in RFID systems, the precise value of N is usually not available. Hence the throughput depends on the estimate of N from tag replies.

During the reading process, the reader can estimate the tag population based on the outcomes of the slots: whether they are empty, singleton or collided. There is usually a misunderstanding that the reader should use several frames of contention slots to estimate the tag population before the real reading process starts. Actually, it is only useful for the earlier RFID systems, which do not support the reservation mechanism. Since the collision of the operation slots would waste more time, the earlier strategies prefer to use a sequence of '*training*' slots (short slots) to estimate the tag population and use it to set the frame size of the *operation slots*. However, in modern RFID systems, a singleton contention slot can reserve an operation slot. Tags can be identified while the reader is doing estimation. Thus the 'training' sequence approach is abandoned. Most algorithms do estimation throughout the reading process.

Based on the estimation methods, algorithms can be divided into: the max-likelihood approach and the probability distribution approach.

3.1 Max-likelihood Approach

Schoute [10] noticed that when N is large and L is suitably chosen (say $L \approx N$), the number of tags attempting each slot has a Poisson distribution with mean 1. The number of collided tags N_C at the end of a frame can be estimated as:

$$N_C = \text{round}(2.39 s_c) \quad (2)$$

where s_c is the number of collided slots in the frame. Therefore his strategy is to set $L = \text{round}(2.39 s_c)$ as the next frame size.

Vogt [11] improved Schoute's strategy by using the statistics of empty slots s_e and singleton slots s_s in addition. Tag population is estimated to be the value N that minimizes the error between the observed values of s_e, s_s, s_c and their expected values using N.

Kodialam [16] proposed an new estimation method based on the Central Limit Theorem. That is when the number of contending tags is large enough, the number of collision slots and empty slots in the current frame should obey the Normal distribution. Thus using his method, one may obtain the estimation accuracy as well as the max-likelihood tag population. But the frame size is also set as $L = E[N]$.

Another example is the Q algorithm in EPCglobal standards [2]. The reader maintains a floating-point variable Q_{fp}. It decreases a typical value C when no tag replies, increases C when multiple tags reply and stays unchanged when only 1 tag replies.[3] The frame size is set to 2^Q, where $Q = \text{round}(Q_{fp})$ and will be canceled whenever $\text{round}(Q_{fp})$ changes. In [13][14], the efficiency of the Q algorithm was obtained with different choices of C and Q_{fp} and some methods to improve efficiency were proposed.

In summary, algorithms of this type compute the maximum-likelihood tag population $\hat{N}$ based on the reading results and set $L = \hat{N}$ as the frame size. This approach is simple but rough, as the expectation only cannot fully describe the variable N.

3.2 Probability Distribution Approach

Floerkemeier [12][15] designed some new strategies based on (1). He assumes that a rough estimation of the target group size is always available in the form of a distribution $\Pr\{N = i\}$ and derives the next frame size as

$$L^* = \left\{ L : \max_{L \in \Upsilon} \sum_{i=0}^{N_{max}} U(N = i, L) \Pr\{N = i\} \right\}, \quad (3)$$

where Υ is the set of possible frame sizes. In every time slot, the reader updates the distribution by Bayesian method and cancels the current frame whenever L^* changes according to (3).

This approach can track the value of N more accurately. Since a random variable N is completely specified by its distribution and Bayesian method ensures no information loss in estimation, the *Population Estimation* part of Floerkemeier's algorithm is undisputable, but the use of (1) in the *Reading Strategy Determination* part is unwise.

3.3 The Need for a New Model

From this review, we can see that previous works focused on the *Population Estimation*, providing different ways to find a more accurate N. For the *Reading Strategy Determination* part, they all use (1) for calculating throughput. As we mentioned before, (1) is obtained from the theory of Random Access (RA) system. In RA system, the frame size is chosen to optimize the instantaneous throughput U. Since a terminal in an RA system would still attempts the channel after a successful transmission, the 'contending group' can be assumed unchange during a long enough period. The long-term throughput of a RA system is therefore equal to the expected instantaneous throughput U calculated by (1). However, in RFID systems, identified tags are silenced by the reader, leading to tag population decrease during the reading process. When the frames are not identical, a concatenation of locally optimal solutions is not globally optimal. As an example, suppose the target group size is distributed as

$$\Pr\{N = i\} = \begin{cases} 0.99 & , \quad i = 0 \\ 0.01 & , \quad i = 10 \end{cases}$$

From (3), the suitable frame size should be $L = 10$, as it can maximize the throughput of the current frame. However, since this group is very likely empty, it is better to use $L = 1$ to check whether it contains tags or not even though the throughput of this checking frame is 0.

[3]In EPCglobe standards, it is recommended that $0.2 \leq C \leq 0.5$ and the initial $Q_{fp} = 4$

4. READING STRATEGY OPTIMIZATION

We now present our method to find the globally optimal frame size. This method can be used for all types of FA algorithms. In this paper, We use the Type 2 algorithm to illustrate.

4.1 The Optimal Reading Strategy

Since canceling a frame is not allowed in Type 2 algorithms, the reading strategies are restricted to the choice of the next frame size. To choose a suitable frame size L, the reader needs the information of the target group size. As discussed in Section 3, this information can be fully described by a probability distribution. In applications, a rough distribution is often available as the reader has information of its previous readings. In the worst case where N is completely unknown, a uniform distribution on $[0, N_{max}]$ can be assumed as we cannot favor any value over the others.

During the reading process, let $Bel(N)$ denote the *belief* of N, or the conditional distribution of N based on all available information [17]. At the end of every frame, the belief can be updated by the Bayesian method [17]. To simplify the notation, let $v_n = Bel(N = n)$ and $\mathbf{v} = (v_0, v_1, \ldots, v_{max})$. Obviously the accuracy of the belief affects the reading efficiency. Let $T(n \,|\, \mathbf{v})$ denote the expected contention time, measured by slots, for these n tags when the current belief is $\mathbf{v}$. Then the expected finishing time is

$$\mathcal{T}(\mathbf{v}) = \sum_{n=0}^{N_{max}} v_n T(n \,|\, \mathbf{v}).$$

Our goal is to find the optimal frame size L^* that can minimize $\mathcal{T}(\mathbf{v})$ for any given distribution $\mathbf{v}$, or

$$L^* = \left\{ L : L \in \Upsilon, \min_{L} \{\mathcal{T}(\mathbf{v})\} \right\}. \tag{4}$$

Note (4) is different from (3) as it is designed to minimize the expected reading time $\mathcal{T}(\mathbf{v})$ instead of the expected instantaneous throughput U. Thus L^* is the globally optimal frame size. To find it, however, requires deriving the function of $\mathcal{T}(\mathbf{v})$ from the reading mechanism of Type 2 algorithms.

In an intelligent system, the optimal decision depends only on the current information, or the belief of all the relevant variables [17]. Applying to RFID systems, the optimal frame size depends only on $Bel(N)$.[4] We let $V_j = Bel(N_j) = (v_0, v_1, \ldots, v_{max})$ denote the state of the reading process at the end of frame j, where N_j is unresolved tag population at the end of frame j. Since identified tags are silenced by the reader, we always have $N_j \geq N_{j+1}$. Further let

$$\mathscr{V} = \left\{ (v_0, v_1, \ldots, v_{max}) \,\middle|\, v_i \geq 0, \sum_{i=0}^{N_{max}} v_i = 1 \right\}$$

denote the set of all possible states. For a group with the initial estimation $P(N)$, let $V_0 = P(N)$ be the initial state and $V_T = (1, 0, 0, \ldots, 0)$ be the terminal state.

THEOREM 1. *Following a distribution-based anti-collision algorithm, the reading process $V_0V_1V_2 \ldots V_T$ is a Markov Chain.*

PROOF. At the end of frame j, let $V_j = (v_0, v_1, \ldots, v_{max}) \in \mathscr{V}$ be the current state. For a distribution-based algorithm, the next frame size l should be fixed given V_j. Let $V_{j+1} = (u_0, u_1, \ldots, u_{max})$ be the belief of tag population at the end of frame $j + 1$. Obviously it depends on the reading results of frame $j + 1$ as well as the previous beliefs.

In frame $j + 1$, let random variable S_0, S_1, S_c denote the number of empty slots, singleton slots and collided slots. It can be proved that the position of the empty slots, singleton slots and collided slots does not matter and only their total numbers affect the belief. Since $S_0 + S_1 + S_c = l$, there are at most $\binom{l}{2} = \frac{1}{2}(l+1)(l+2)$ different outcomes. Thus for a given frame size, there are at most $\frac{1}{2}(l+1)(l+2)$ different choices of V_{j+1} that satisfy $\Pr(V_{j+1} \,|\, V_j) > 0$.

Analogous to the urn problem [18], the probability that s_1 urns contain only 1 ball, s_c urns contain more than 1 balls and the others are empty can be obtained as:

$$\Pr\{S_c = s_c, S_1 = s_1 \mid N_j = n, L = l\}$$

$$= \binom{l}{s_0, s_1, s_c} \frac{n!}{(n-s_1)!\,l^n} \sum_{\substack{m_1, m_2, \ldots, m_{sc} \geq 2, \\ m_1 + m_2 + \cdots + m_{sc} = n - s_1}} \binom{n - s_1}{m_1, m_2, \ldots, m_{sc}}, \tag{5}$$

where $m_1, m_2, \ldots, m_{sc}$ denote the number of tags in each of the s_c collided slots. Further, we can substitute the belief of N_j to obtain

$$\begin{aligned} &\Pr\{S_c = s_c, S_1 = s_1 \mid L = l\} \\ &= \sum_{n=0}^{N_{max}} v_n \ \Pr\{S_c = s_c, S_1 = s_1 \mid N_j = n, L = l\}. \end{aligned} \tag{6}$$

At the end of frame $j + 1$, we can obtain the values of s_0, s_1 and s_c. By Bayes formula, the posterior distribution of N_j can be updated as:

$$\begin{aligned} v_i' &= \Pr\{N_j = i \mid S_c = s_c, S_1 = s_1, L = l\} \\ &= \frac{\Pr\{S_c = s_c, S_1 = s_1 \mid N_j = i, L = l\}\ v_i}{\Pr\{S_c = s_c, S_1 = s_1 \mid L = l\}} \end{aligned} \tag{7}$$

As tags in the singleton slots are successfully identified and silenced, we have $N_{j+1} = N_j - s_1$ with distribution given as

$$\begin{aligned} u_i &= \Pr\{N_{j+1} = i \mid S_c = s_c, S_1 = s_1, L = l\} \\ &= \Pr\{N_j = i + s_1 \mid S_c = s_c, S_1 = s_1, L = l\} \\ &= v_{i+s_1}', \qquad i = 0, 1, 2, \ldots \end{aligned} \tag{8}$$

Since the transition probability from state $V_j = (v_0, v_1, \ldots, v_{max})$ to $V_{j+1} = (u_0, u_1, \ldots, u_{max})$ is just $\Pr\{S_c = s_c, S_1 = s_1 \mid L = l\}$, which depends only on V_j and V_{j+1}, the states $V_0V_1V_2 \ldots V_T$ forms a Markov Chain. □

[4]Note it is important to differentiate the 'unconditional optimal' and the 'optimal based on current belief'. As an example, suppose $N = 10$, but our current belief is $Bel(N = 9) = 1$. Then the 'unconditional optimal' frame size is 10, but the 'optimal' frame size based on the current knowledge is 9. Since the unconditional optimal frame size is not available until the reading process is finished, in this paper we only consider the optimal one based on current belief.

For a given state V_j, let $\mathcal{V}(V_j, l) \subset \mathscr{V}$ denote the set of possible V_{j+1}, or

$$\mathcal{V}(V_j, l) = \Big\{V_{j+1} \mid \Pr(V_{j+1} \mid V_j) > 0\Big\}.$$

As proved in Theorem 1, $|V(V_j, l)| \le \frac{1}{2}(l+1)(l+2)$. Let $\mathcal{T}_O(V_j)$ denote the first passage time from V_j to V_T using the optimal reading strategy. From the theory of Markov Chain [19], we have

$$\begin{aligned}\mathcal{T}_O(V_j) &= L^* + \sum_{V_{j+1}\in\mathcal{V}(V_j,\,L^*)} \Pr(V_{j+1} \mid V_j)\mathcal{T}_O(V_{j+1}) \\ &= \min_l \left\{ l + \sum_{V_{j+1}\in\mathcal{V}(V_j,\,l)} \Pr(V_{j+1} \mid V_j)\mathcal{T}_O(V_{j+1}) \right\} \quad (9)\end{aligned}$$

In (9), the set $\mathcal{V}(V_j, l)$ and the transition probability $\Pr(V_{j+1} \mid V_j)$ are given by (8) and (6). Theoretically speaking, it can be solved to obtain the optimal frame size L^*. In the following, we show how to solve (9) analytically and numerically.

4.2 The Analytical Solution of $\mathbf{L^*}$

In this section, we show how to solve (9) by some examples. Since (9) is a recursive function, we begin from small N_{max} cases

4.2.1 *Case 1:* $N_{max} = 2$

In this case, the tag population can only be 0, 1 or 2. Given $V_j = \mathbf{v} = (v_0, v_1, v_2)$, the probability of different outcomes of frame $j+1$ can be obtained from (6) as:

$$\begin{aligned}&\Pr\{S_c = 0, S_1 = 0 | L = l\} = v_0;\\ &\Pr\{S_c = 0, S_1 = 1 | L = l\} = v_1;\\ &\Pr\{S_c = 0, S_1 = 2 | L = l\} = \frac{l-1}{l}v_2;\\ &\Pr\{S_c = 1, S_1 = 0 | L = l\} = \frac{1}{l}v_2.\end{aligned}$$

From (7) and (8), we get the distribution of N_{j+1} as

$$\begin{aligned}&\Pr\{N_{j+1} = 0 | S_c = 0, S_1 = s_1, L = l\} = 1;\\ &\Pr\{N_{j+1} = 2 | S_c = 1, S_1 = 0,\ L = l\} = 1.\end{aligned}$$

Thus V_{j+1} has only two possible choices as

$$\mathcal{V}(V_j, l) = \Big\{(1,0,0),\ (0,0,1)\Big\}$$

and the transition probability is

$$\begin{aligned}\Pr\Big((1,0,0)\,\Big|\,(v_0,v_1,v_2)\Big) &= \frac{l-1}{l}v_2 + v_1 + v_0\\ \Pr\Big((0,0,1)\,\Big|\,(v_0,v_1,v_2)\Big) &= \frac{1}{l}v_2\end{aligned}$$

Substituting them into (9), we have

$$\begin{aligned}\mathcal{T}_O(\mathbf{v}) &= \min_l \Big\{ l + \frac{1}{l}v_2\ \mathcal{T}_O\Big((0,0,1)\Big)\Big\}\\ &= \min_l \Big\{ l + \frac{4v_2}{l}\Big\}, \qquad (10)\end{aligned}$$

where $T_O\Big((0,0,1)\Big) = 4$ is the expected reading time for a group with exactly 2 tags, which can be obtained in *'Case 3'*. So we have

$$L^* = \begin{cases} 1 & , \quad v_2 < \frac{1}{2}\\ 2 & , \quad v_2 \ge \frac{1}{2}\end{cases}$$

$$\mathcal{T}_O(\mathbf{v}) = \begin{cases} 1 + 4v_2 & , \quad v_2 < \frac{1}{2}\\ 2 + 2v_2 & , \quad v_2 \ge \frac{1}{2}\end{cases}$$

To compare, we derived the frame size and average reading time of Floerkemeier's strategy from (3) as

$$L = \begin{cases} 1 & , \quad v_1 > v_2\\ 2 & , \quad v_1 \le v_2\end{cases}$$

$$\mathcal{T}_f(\mathbf{v}) = \begin{cases} 1 + 4v_2 & , \quad v_1 > v_2\\ 2 + 2v_2 & , \quad v_1 \le v_2\end{cases}$$

Therefore, when the distribution $\mathbf{v}$ satisfies $v_1 < v_2 < 0.5$, Floerkemeier's strategy is not optimal.

4.2.2 *Case 2:* $N_{max} = 3$

Next, we move on to $N_{max} = 3$ and derive the recursive function as

$$\mathcal{T}_O(\mathbf{v}) = \min_l \left\{ l + \frac{12v_3(l-1)}{l^2} + (\frac{v_2}{l} + \frac{v_3}{l^2})\mathcal{T}_O(\mathbf{u})\right\}, \quad (11)$$

where $\mathbf{u}$ is the distribution of N_{j+1} on condition that $S_c = 1$ and $S_1 = 0$ in frame $j+1$, or

$$\mathbf{u} = (u_0, u_1, u_2, u_3) = \left(0, 0, \frac{lv_2}{v_3 + lv_2}, \frac{v_3}{v_3 + lv_2}\right).$$

Solving (11), the optimal reading strategy can be similarly obtained as:

$$L^* = \begin{cases} 1 & , \quad \mathbf{v} \in \Phi_1\\ 2 & , \quad \mathbf{v} \in \Phi_2\\ 3 & , \quad \mathbf{v} \in \Phi_3\end{cases}$$

$$\mathcal{T}_O(\mathbf{v}) = \begin{cases} 1 + (v_2+v_3)T_x\left(\frac{v_3}{v_2+v_3}\right) & , \quad \mathbf{v} \in \Phi_1\\ 2 + 3v_3 + \frac{1}{4}(2v_2+v_3)T_x\left(\frac{v_3}{2v_2+v_3}\right) & , \quad \mathbf{v} \in \Phi_2\\ 3 + \frac{8}{3}v_3 + \frac{1}{9}(3v_2+v_3)T_x\left(\frac{v_3}{3v_2+v_3}\right) & , \quad \mathbf{v} \in \Phi_3\end{cases}$$

where T_x is a recursive function as

$$T_x(\alpha) = \begin{cases} \frac{8}{3}\alpha + 4 & , \quad 0 < \alpha \le \frac{9}{16}\\ 3 + \frac{8}{3}\alpha + \frac{1}{9}(3-2\alpha)T_x\left(\frac{\alpha}{3-2\alpha}\right) & , \quad \frac{9}{16} < \alpha < 1\end{cases}$$

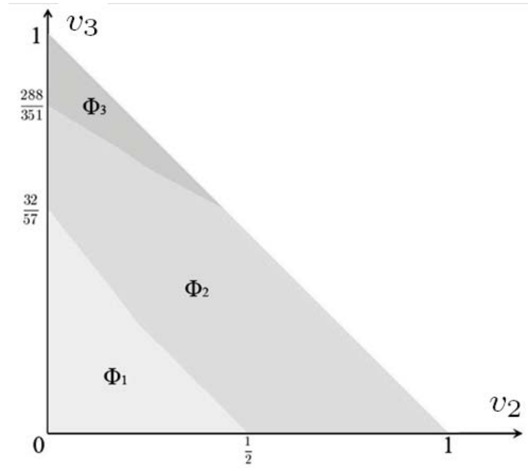

Figure 3: The solution set for $N_{max} = 3$ case

while Φ_1, Φ_2 and Φ_3 are distribution regions specified in Figure 3. It shows that the choice of L is determined by v_2 and v_3 instead of $E[N_A]$. As an example, suppose Group X has tag population distribution as (0, 0.45, 0, 0.55) and Group Y (0.4, 0, 0.5, 0.1). Their expected group size are calculated to be $E[N_X] = 2.1$ and $E[N_Y] = 1.3$, but the optimal frame size for Group X is 1 while that for Group Y is 2.

For $N_{max} > 3$ cases, the optimal strategies can be obtained similarly. But the computation becomes more complex.

4.2.3 *Case 3: Tag Population N known*

Let $\mathcal{T}_N(n)$ denote the average reading time for a group with tag population known as n, or $\mathcal{T}_N(n) = \mathcal{T}_O(\mathbf{v})$ when $v_n = 1$. After one frame of reading, the population decreases to $n' = n - s_1$, where s_1 is the number of singleton slots. So the remaining reading time is $\mathcal{T}_N(n - s_1)$. Substituting $\mathcal{T}_N(n)$ and $\mathcal{T}_N(n - s_1)$ into (9), we have

$$\mathcal{T}_N(n) = \min_l \left\{ l + \sum_{s_1=0}^{l} \sum_{s_c=0}^{l-s_1} \Pr\{S_c = s_c, S_1 = s_1 \mid l\} * \mathcal{T}_N(n - s_1) \right\}. \quad (12)$$

The explicit form of the system equation after rearranging is

$$\mathcal{T}_N(n) = \min_l \left\{ \frac{l + \sum_{s_1=1}^{l} \sum_{s_c=0}^{l-s_1} \Pr\{S_c = s_c, S_1 = s_1 \mid l\} \mathcal{T}_N(n - s_1)}{1 - \sum_{s_c=1}^{l} \Pr\{S_c = s_c, S_1 = 0 \mid l\}} \right\}. \quad (13)$$

Solving (13), the optimal frame size is obtained as $L^* = n$. The algorithm efficiency $\eta = \frac{n}{\mathcal{T}_N(n)}$ is shown in Figure 4. We can see that η decreases with n and approaches e^{-1}, which coincides with the efficiency upper bound of Random Access systems [10].

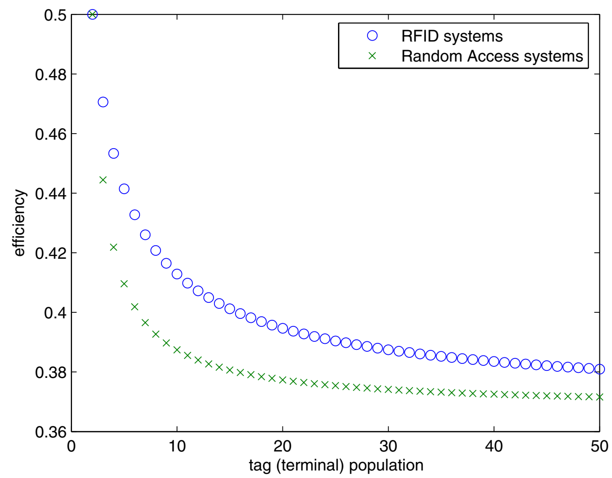

Figure 4: The efficiency of the Optimal Type 2 FA algorithm when tag population N is known

4.3 The Numerical Solution of L^*

Here, we introduce a method to compute the values of L^* by running a recursive program.

From the last section, we know for a group with known population k, the optimal frame size is $L^* = k$ and the average reading time is $\mathcal{T}_N(k)$. It is easy to imagine for a group with tag population very likely to be k, the optimal frame size should also be $L^* = k$. In other words, if the distribution of N satisfying:

1. $v_k > v_i$, where $i \neq k$ and $0 \leq i \leq N_{max}$, $1 \leq k \leq N_{max}$;
2. $\mathrm{Var}(\mathbf{v}) < \delta$, where δ is a small enough value,

the optimal frame size is still $L^* = k$. If it does contain k tags, obviously, the average reading time approaches $\mathcal{T}_N(k)$ when $\delta \to 0$, or $\lim_{\delta \to 0} T(k \mid \mathbf{v}) = \mathcal{T}_N(k)$. But if it contains j tags, where $j \neq k$, we let $\lim_{\delta \to 0} T(j \mid \mathbf{v}) = \mathcal{T}_N^{(k-j)}(j)$, or $\mathcal{T}_N^{(k-j)}(j)$ is the average reading time for j tags starting with the wrong information that $N \to k$. Averaging all these cases, we have

$$\lim_{\delta \to 0} \mathcal{T}(\mathbf{v}) = \sum_{j=0}^{N_{max}} \lim_{\delta \to 0} v_j T(j \mid \mathbf{v}) = \sum_{j=0}^{N_{max}} v_j \mathcal{T}_N^{(k-j)}(j), \quad (14)$$

where $\mathcal{T}_N^{(0)}(j) = \mathcal{T}_N(j)$ for $j = k$. Similar to the derivation in Case 3, the formula for $\mathcal{T}_N^{(m)}$ can be obtained as: for $m < 0$,

$$\mathcal{T}_N^{(m)}(n) = \left\{ L + \sum_{S_1=1}^{L} \sum_{S_c=0}^{X} \Pr\{S_c, S_1 | L\} * \mathcal{T}_N^{(m)}(n - S_1) + \sum_{k=0}^{-m-1} \sum_{S_1=0}^{n+m} \Pr\{S_c = \frac{n - S_1 - k}{2}, S_1 | L\} * \mathcal{T}_N^{(-k)}(n - S_1) \right\} * \left\{ 1 - \sum_{s_c=1}^{X} \Pr\{S_c = s_c, S_1 = 0 \mid L\} \right\}^{-1}, \quad (15)$$

and for $m \geq 0$,

$$\mathcal{T}_N^{(m)}(n) = \frac{L + \sum_{S_1=1}^{L} \sum_{S_c=0}^{L-S_1} \Pr\{S_c, S_1 | L\} * \mathcal{T}_N^{(m)}(n - S_1)}{1 - \sum_{S_c=1}^{L} \Pr\{S_c, S_1 = 0 | L\}}, \quad (16)$$

where $L = m + n$ and $X = \min(L - S_1, \frac{n - S_1}{2})$. Limited by space, we skip the mathematical details.

With the above formulas, the computer program is designed as follows[5]:

[5]This program is designed for $S_L = \{1, 2, 3, 4, \ldots\}$ case. If L is limited to the powers of 2, the program is simpler.

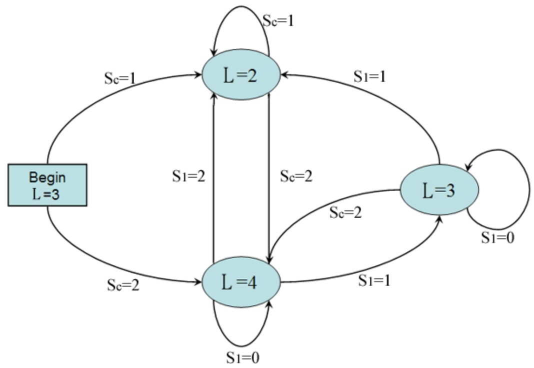

Figure 5: State machine (N uniform distributed in $[1, 5]$)

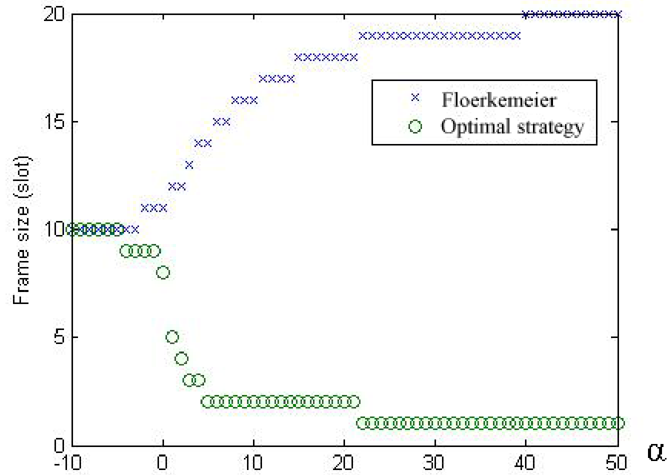

Figure 6: The frame size for different value of α

```
=================================
Function [T_v, L_v]=OptimalAloha(v)          // realize(9)

if (var(v) < δ)                      // the stopping condition
    T_v = T(v);                                // using (14)
    L_v = E[N];
    return;
end if
l =round(E[N]); T_v = ∞;
for L_t = l − Δl : l + Δl            // numerically find L*
    temp =0;
    for s_1 = 0 : L_t
        for s_c = 1 : L_t − s_1
            Calculate Pr{s_c, s_1|L_t} from (6)
            Calculate b from (7) and (8);
            [T_u, L_u]=OptimalAloha(u);
            temp=temp+ Pr{s_c, s_1|L_t}T_u;
        end for
    end for
    if(T_v >temp)
        T_v=temp; L_v = L_t;
    end if
end for
=================================
```

The input is a vector $\mathbf{v}$ representing the distribution. The program first checks its variance. If it is smaller than a threshold δ (our experiment shows $\delta \approx 0.4$ is enough), the optimal frame size is $E[N]$ and the average reading time is calculated from (14); if not, it is numerically resolved. Since the variance of the input distribution decreases as the program is recursively used, the stopping condition will be fulfilled after several loops.

For implementation, these results can be precalculated and stored in database. There is no need to do any computation during the reading process. For example, when tag population is uniformly distributed from 1 to 5, the optimal strategy is shown in Figure 5 as a state machine.

5. APPLICATION EXAMPLES

In a modern supermarket where all the merchandize are tagged, customers just need to walk their carts through a door for all items to be identified. Let N denote the number of items in one customer's cart. Although the precise value of N is usually unknown, a distribution of N is often available from the past sales statistics.

For a given distribution, most of the previous algorithms can give the optimal frame size when the variance of N is small. But when the variance of N is large, the frame size is often inappropriately chosen. Here we use a simple example to show that the optimal algorithm we proposed still works efficiently for large-variance samples.

Consider an express check-out supermarket counter where each customer is allowed to checkout no more than 20 items, or $N \le 20$. To illustrate the effect of population variation on reading performance, we set $E[N] = 10$ and change the variance. $\mathrm{Var}[N] = 0$ when all customers buy exactly 10 items each. $\mathrm{Var}[N]$ is maximized when half of them buy 20 items while the other half buy nothing. In our experiment, we choose:

$$Pr\{N = n\} = \begin{cases} \frac{1}{Z}\left(|n-10|+10\right)^{\alpha} & , \quad 0 \le n \le 20, \\ 0 & , \quad \text{others} \end{cases}$$

where Z is the normalization constant and α is a variable. The variance of N_A increase with α while $E[N] = 10$ is independent of α. Specifically, when $\alpha \to -\infty$, $\mathrm{Var}(N)$ approaches 0; when $\alpha = 0$, N is uniformly distributed in $[0, 20]$; and when $\alpha \to \infty$, $\mathrm{Var}(N)$ is maximum. (Note that this distribution is chosen for simplicity. Other distributions we tried give similar results.) Following Schoute's and Vogt's strategies ($L = E[N]$), the 'suitable' frame size is just 10 regardless the choice of α. [6] For distribution-sensitive algorithms (Floerkemeier's algorithm and the Optimal Type 2 algorithm), the choices of frame size are listed in Figure 6. The frame size of both algorithms start with 10 for small variance cases, but diverge to 20 and 1 respectively as the

[6] We do not simulate Kodialam's algorithm [16], because in his algorithm the reader uses several *'training'* frames to estimate tag population before the real reading process starts. As mentioned in section 3, this is not efficient for modern RFID systems. Thus its reading time would be much longer compared other algorithms.

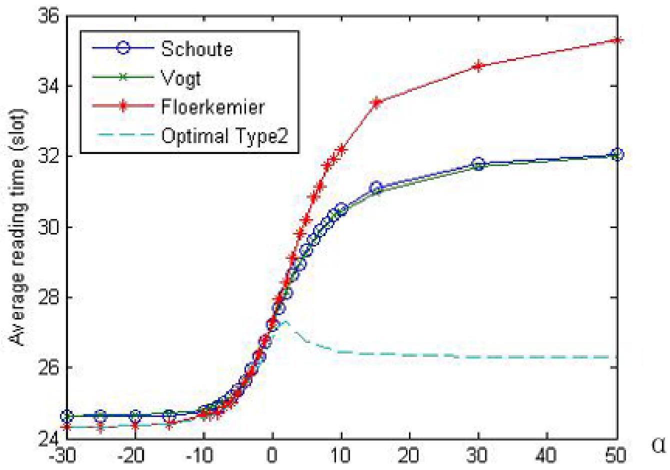

Figure 7: The average reading time from simulation

variance increases. Figure 7 shows the average reading time of four strategies from computer simulation (the average of 1 million samples for each point) as a function of Var(N). We observe:

- The performance of Schoute's and Vogt's algorithms are barely distinguishable.
- Floerkemeier's algorithm is marginally better than Schoute's when $\alpha < -2$, or Var(N) < 24 but poorer when above.
- The Optimal Type 2 algorithm is the best for all values of α and the performance gain increases with the variance of tag population;
- The minimum performance gain is obtained at $\alpha \rightarrow -\infty$, or Var($N$) $= 0$, where the average reading time of the Optimal Type 2 algorithm is 24.2 slots, the same as that of Floerkemeier's algorithm.
- The maximum performance gain is obtained at $\alpha \rightarrow \infty$, where the average reading time of the Optimal Type 2 algorithm is reduced from 35.3 slots to 26.3 slots when compared to Floerkemeier's algorithm.

The Optimal Type 2 algorithm is only marginally better for small variance cases, because it is usually easier to find the suitable frame size when the target group size does not change dramatically. However, as the variance increases, traditional algorithms fail to make suitable choices while the optimal algorithm can still work efficiently. Since the samples to identify usually have large variance in real application, the improvement is considerable.

6. SUMMARY

In this paper, we proposed a new optimizing method for Framed Aloha based RFID anti-collision algorithms. It makes globally optimal decisions based on the current information. The optimal parameters can be obtained by running an iterative program and adopted to different applications. Simulation results show that significant improvement was obtained, especially when the variance of tag population is large.

7. ACKNOWLEDGMENTS

This work was partially supported by the Competitive Earmarked Research Grant (Project Number CUHK418507) established under the University Grant Committee of the Hong Kong Special Administrative Region, China.

8. REFERENCES

[1] K. Finkenzeller. RFID handbook - Second Edition. JOHN WILEY & SONS, 2003.

[2] EPCglobal. EPCglobal Class 1 Generation 2 UHF Air Interface Protocol Standard Version 1.0.9, 2005. http://www.epcglobalinc.org/.

[3] International Organization for Standardization. Information technology - RFID for item management - Part 6: Parameters for air interface communications at 860 MHz to 960 MHz, 2004.

[4] Philips Semiconductor. I-CODE1 Label ICs Protocol Air Interface Datasheet, January 2005.

[5] J. Myung, W. Lee, J. Srivastatva, T.K. Shih. Tag-Splitting: Adaptive Collision Arbitration Protocols for RFID Tag Identification. *IEEE Transactions on Parallel and Disitributed Systems*, Vol 18, No.6, June 2007.

[6] K.W. Chiang, C.Q. Hua and T.S.P. Yum. Prefix-Randomized Query-Tree Protocol for RFID Systems. *IEEE ICC*, 2006

[7] P. Popovski, F.H.P. Fitzek, R. Prasad. Batch Conflict Resolution Algorithm with Progressively Accurate Multiplicity Estimation. *ACM DIALM-POMC* Oct. 2004.

[8] J.Park, M.Y. Chung and T.J. Lee. Identification of RFID Tags in Framed-Slotted ALOHA with Tag Estimation and Binary Splitting, *Communications and Electronics*, 2006

[9] J. Mosely, P.A. Humblet. A Class of Efficient Contention Resolution Algorithms for Multiple Access Channels. *IEEE Transactions of communications, Vol. Com-33, No.2*, Feb. 1985.

[10] F.C.Schoute. Dynamic Frame Length ALOHA. *IEEE Transactions on Communications*, COM-31(4):565-568, Apr 1983.

[11] H. Vogt. Efficient Object Identification with Passive RFID Tags. *First International Conference, PERVASIVE 2002, volume 2414 of Lecture Notes in Computer Science (LNCS)*, pages 98-113, Zurich, Switzerland, August 2002. Springer-Verlag.

[12] C. Floerkemeier. Transmission control scheme for RFID object identification. *Proceedings of the Pervasive Wireless Networking Workshop at IEE PERCOM 2006*, Pisa, Italy, 2006

[13] B. Zhen, M. Kobayashi, M. Shimizu. Framed Aloha for Multiple RFID objects Identification. *IEICE Transactions of comminication. Vol.E88-B, No.3*, March 2005

[14] M. Buettner, D. Wetherall. An Empirical Study of UHF RFID Performance. *MobiCom*, San Francisco, California, USA, 2008

[15] C. Floerkemeier. Bayesian Transmission strategy for Framed ALOHA Based RFID Protocols. *IEEE International Conference on RFID*, Gaylord Texan Resort, Grapevine, TX, USA, March, 2007

[16] M. Kodialam, T. Nandagopal. Fast and Reliable Estimation Schemes in RFID Systems. *MobiCom*, Los Angeles, California, USA, Sepetember, 2006

[17] J. Pearl. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference, Morgan Kaufmann Publishers.

[18] W. Feller. An Introduction to Probability Theory and Its Applications, Second edtion, Vol I, John Wiley.

[19] S.M. Ross. Introduction to Probability Moldels, Seventh Edition, Chapter 6, Academic Press.